

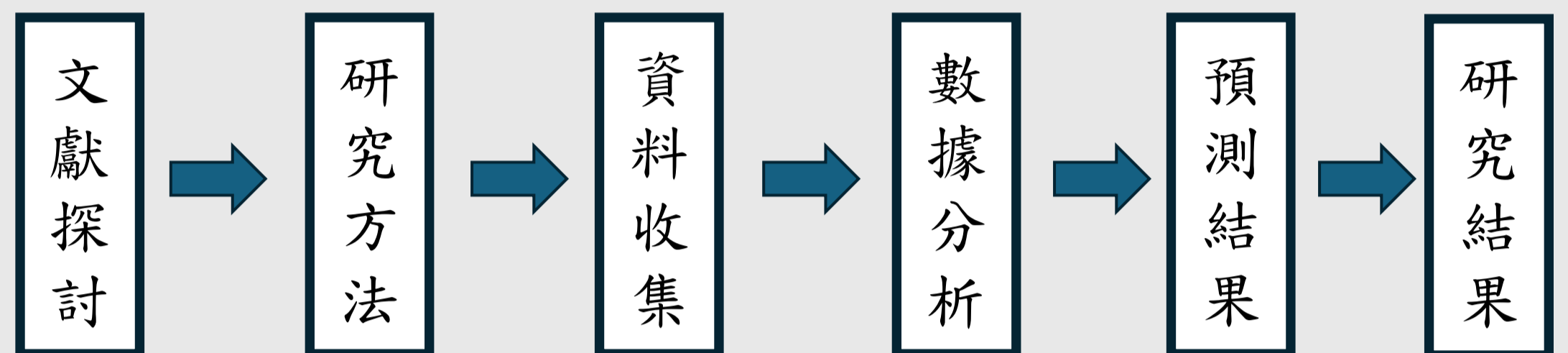


研究動機與目的

台灣機車市場高度成熟，KYMCO 與 SYM 長期占據市占率前二。然而，過去市佔率分析相關討論多以主觀觀察為主，缺乏以數據為基礎的系統性分析。隨著資料分析與機器學習技術的進展，本研究希望透過更客觀的量化方法，探討兩大品牌的市占率趨勢及其主要影響因素，以協助消費者與企業理解市場變化並做出更有效的策略決策。提出對 SYM 與 KYMCO 未來市場策略之實證參考。

研究範圍及架構

- 1. 中華民國交通部公路局中2018至2025年間SYM 與 KYMCO機車之數據。
2. 變數數據為Y=市佔率(%)。
3. X=銷售量、價格、CP值主打強度、各年度車款數量。



研究方法

(1) 移動平均

本研究利用移動平均法所預估出的價格、CP值主打強度、各年度車款數量的未來值，作為未來市佔率的變數預測，以為線性回歸跟隨機森林後續的使用。

(2) 線性趨勢

我們單獨對銷售量使用線性趨勢，因為銷售量隨時間的變化通常具有「穩定的長期趨勢」，而且資料的波動呈現近似線性的上升或下降，所以用線性趨勢能捕捉到整體方向，在缺少更複雜外部變數時，是最合理且常用的預測方法。

(3) 線性回歸

本研究採用 Python 之 scikit-learn 套件中的 sklearn.linear\_model.LinearRegression 函式建立線性回歸模型，旨在評估影響 SYM 與 KYMCO 月市占率變動的主要因素。模型中選取之自變數包含各月份的平均價格、月銷售數量、CP 值主打強度指標，以及各年度所推出之車款總數等。

(4) 隨機森林

本研究同時採用 Python 之 scikit-learn 套件中的 sklearn.ensemble.RandomForestRegressor 演算法建立隨機森林預測模型，隨機森林屬於集成式 (Ensemble) 機器學習模型，其透過建立多棵決策樹並進行整合，能有效捕捉資料中的非線性關係及複雜交互作用。

研究結果

1. 建立線性回歸模型的特徵係數

```
python
coef_kym = {
  "銷售量": 2.1831866422063183e-06,
  "avg_price": 0.0085816044343340784,
  "cp_value_score": 0.16942888870722613,
  "各年度車款數量": -0.0031208979756167394,
  "intercept": 0.11705204325847263
}
coef_sym = {
  "銷售量": 3.39407672891205e-06,
  "avg_price": 0.0452732982435158,
  "cp_value_score": 0.23876213407536842,
  "各年度車款數量": -0.04053850357869169,
  "intercept": 0.23262605297077693
}
```

兩品牌的銷售量皆是正向影響力最大的特徵，增加它們會顯著的提升市占率。

2. 線性回歸

(1) 使用LinearRegression函式：建立線性回歸模型。

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score, mean_squared_error
```

(2) 建立線性回歸模型。

```
lr = LinearRegression()
lr.fit(X_train, y_train)
y_pred_lr = lr.predict(X_test)
```

(3) 預測結果

```
=== 線性回歸模型結果 ===
SYM (線性回歸) -> MSE: 0.001153, R²: 0.7957
KYMCO (線性回歸) -> MSE: 0.001394, R²: 0.5016
```

4. 未來六個月的預測結果

(1) 使用隨機森林的結果

Table with 4 columns: 品牌, 日期, RF\_市佔率. Rows show predictions for SYM and KYMCO from 2025-07 to 2025-12.

(2) 使用線性回歸的結果

Table with 4 columns: 品牌, 日期, LR\_市佔率. Rows show predictions for SYM and KYMCO from 2025-07 to 2025-12.

3. 隨機森林

(2) 建立隨機森林模型。

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import r2_score, mean_squared_error
```

```
rf = RandomForestRegressor(random_state=42, n_estimators=200)
rf.fit(X_train, y_train)
y_pred_rf = rf.predict(X_test)
```

(3) 預測結果

```
=== 隨機森林模型結果 ===
SYM (隨機森林) -> MSE: 0.000286, R²: 0.9493
KYMCO (隨機森林) -> MSE: 0.001373, R²: 0.5091
```

(1)使用函式

sklearnRandomForestRegressor建立隨機森林模型。

結論

隨機森林模型在處理較複雜、非線性的市場行為時效果更佳。

三陽 SYM 預測市佔率：44%
光陽 KYMCO 預測市佔率：約 27%

未來方向

- 1.加強模型準確性 3.跨領域合作與應用
2.增加數據多樣性 4.擴增研究範疇及應用場景